

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

MEASURING GOVERNMENT EDUCATION OUTPUT IN THE NATIONAL ACCOUNTS

**An overview of several methods developed as part of
the Atkinson Review**

department for

education and skills

creating opportunity, releasing potential, achieving excellence

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

EXECUTIVE SUMMARY

Overview

1. The Department took part in the independent review of government output and productivity chaired by Sir Tony Atkinson which sought better output measures for education for the National Accounts. Specifically, we developed a number of ways of better measuring the quality of education outputs. We subjected these measures to a peer review which included independent academics. As a result, we will recommend to the Office for National Statistics (ONS) a quality measure based on GCSE attainment thresholds, and an overall output measure adjusted by the trend rate of real earnings growth.

Current measure of education output in National Accounts

2. Output in the publicly-funded compulsory education sector is currently measured using pupil numbers, with a fixed quality adjustment based on the historic trend in GCSE results. The adjustment is based on data from the mid-1990s, and reflects ONS' assumption that education is a multiplicative – rather than additive – process. The adjustment increases output by 0.25% per year, which is around a tenth of the increase in raw GCSE results.
3. Atkinson recommended investigating improving the output measure by: (a) adjusting pupil numbers by attendance rates (which was implemented for Blue Book 2005); (b) improving or updating the quality adjustment and (c) devising a means of reflecting the changing value of education output.

Alternative measures investigated

4. The Peer Reviewer Group considered that – based on how children develop vocabulary – an additive model of the education process seems more realistic than a multiplicative model. The review looked at three ways of capturing changes in the quality of education output, applied within an additive model to derive alternative education output measures:
 - results of Ofsted school inspections;
 - Improvements in GCSE attainment; and
 - progress of cohorts of pupils through each of the Key Stages.

Ofsted effectiveness measures

5. In principle, the “overall effectiveness” judgements from Ofsted's school inspections could be used to derive a quality index. They have the advantage of covering much more of what goes on in school than simply attainment. However, there are serious drawbacks for National Accounting purposes: the framework for assessment has changed over

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

the years; and we need to weight together the Ofsted results in a rather arbitrary way, as the grading scale is not necessarily linear – for example, we would have to assume that a school given an inspection grade of 2 is assumed to be twice as effective as one given a 4, but is this realistic? In 2005, the scale will also be changed from a 7 point to a 4 point scale. And finally, the Ofsted system overall effectiveness judgements only exist from 2000 – not a long enough time-series for National Accounts purposes. Nevertheless, the peer review concluded that Ofsted measures could provide a complementary view of output from the education system to that provided purely by qualifications-based measures.

Improvements in GCSE attainment

6. GCSEs are taken at the end of compulsory phase of education and capture knowledge accumulated during the pupils' educational career to that point. Such measures have the advantage that we have data on them over a long run of years. They are also readily-understood, transparent and easily-interpretable, and arguably, they are directly relevant to what employers value about the education system – i.e. qualifications. And it is relatively simple to develop measures of the labour market value of education output from evidence on the link between qualifications and wages.
7. But the measures have limitations because they use the attainment of only one cohort of pupils in the education system in any given year (i.e. those sitting GCSEs). Their attainment is taken to be a proxy for improvements in the quality of the education system of all pupils across all year-groups in that year, whereas in reality it reflects quality improvements in the education system for one year-group over the preceding 11 years.
8. Nevertheless, despite these, and some other more technical limitations, the peer review favoured the GCSE threshold measure - on the grounds of transparency and the clear potential for developing labour market valuations for education output.

Cohort progress measures (CPM)

9. Ideally, we would like a measure that captured the progress made by all pupils in education in any given year – and only the progress made in that year. Cohort progress measures - explained more fully in the main paper - are designed to get closer to that ideal.
10. Imagine a table in which the columns represent the progress made by a single year-group and the rows represent calendar years. The output of the system in each calendar year is simply the pupil-weighted sum of cells across a row – the annual progress of all 11 cohorts. The output of the system related to each cohort of pupils is the sum of cells down

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

each diagonal, with each diagonal representing a different cohort. So, the GCSE measure is simply equivalent to the sum of one of the diagonals – that representing the cohort sitting their GCSEs in a particular year. Thus, if there was a fall in progress for some cohorts, that took place less than 11 years ago, the cohort progress measure will register a reduction in output, although on the GCSE measure, output will still rise for some years.

11. There are limitations with the cohort progress method. Since children are only assessed at the end of each Key Stage, and a Key Stage can last from 2 to 4 years, some degree of interpolation between years is required to make an estimate of progress in a single year. It was the strictness and number of assumptions that needed to be made to derive CPMs that swung the peer reviewers' judgements about which type of measure was preferable.

Valuing the output of compulsory education

12. Measuring output as a function of the number of pupils, as all these measures do, results in the problem that, when real teachers' real salaries are rising, education will appear to suffer from a continuous decline in productivity (unless there is also growth in pupil numbers and pupil-teacher ratios). But teachers' salaries will have risen because of real growth in productivity in the private marketed sector and because education has to compete with the marketed sector for labour. Tony Atkinson concluded that therefore the value of public sector output should be seen as growing in line with marketed sector output: a qualified person produced in 10 years time has more productive potential than one produced now because of future technological progress. The peer reviewers agreed with Tony Atkinson's suggestion that we assume the value of education output rises by 1.5% per year, reflecting the trend rate of growth in real earnings.

Relationship between output and productivity

13. It follows that – before making allowance for quality improvements –if the real value of education output rises by 1.5% a year, and the real cost of teachers' salaries rises by 2.5% a year, then educational productivity is falling because the real value of outputs is rising at a slower rate than the real cost of inputs. But if quality is rising by more than 1% a year (in this example), productivity will be improving. Atkinson was very clear that the measure of public sector output should be adjusted for quality.
14. Even if education output is rising, eventually diminishing returns may set in. That is to say, there will come a point where the productivity of additional resources put into the sector may be below the average productivity of resources already in it. When this happens, average productivity will start to fall. But this is not a reason to stop expanding

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

resources in the education sector: investment continues to be justified so long as the value of the additional benefits from the investment exceeds its additional costs.

Illustration of method for measuring output recommended to ONS.

15. The table below illustrates the derivation of our recommended measure of education output. The figures are not exactly as they would appear in National Accounts, since they refer only to England and exclude nursery and school 6th-form education.

Derivation of the education output index for maintained primary and secondary education based on revised GCSE threshold measure (England only)							
	1998	1999	2000	2001	2002	2003	2004
Raw pupil numbers ('000 pupils)	6450	6495	6535	6557	6542	6541	6510
1) Index - raw pupil numbers	98.4	99.1	99.7	100.0	99.8	99.8	99.3
Pupil numbers minus absence	5968	6025	6077	6068	6071	6084	6072
2) Volume Index	98.4	99.3	100.2	100.0	100.1	100.3	100.1
% achieving 5 GCSEs at A*-C	43.8%	45.7%	47.1%	47.9%	49.5%	51.1%	52.0%
Quality Index	91.4	95.4	98.3	100.0	103.3	106.7	108.6
3) Output Index	89.9	94.7	98.5	100.0	103.4	107.0	108.6
4) Reflecting real value	86.0	92.0	97.0	100.0	104.9	110.2	113.6

Next steps

16. ONS will discuss the new measure within ONS and through wider debate. ONS will decide whether to submit our proposed measure to its National Accounts Group (NAG) for review. If the NAG accepts our proposals, they may include it in the National Accounts for the future. Meanwhile, data would be gathered and examined through productivity articles, and possibly the construction of an experimental series. ONS will also publish an article on productivity in education during October which draws extensively on our work. Their article will not recommend any particular measure of output, but will consider education productivity from a variety of perspectives.

ATKINSON REVIEW AND THE NATIONAL ACCOUNTS

Atkinson Review

1.1 In December 2003, the National Statistician commissioned Sir Tony Atkinson to carry out an independent review of the measurement of government output and productivity. The aim of the review was “to advance methodologies for the measurement of government output, productivity and associated price indices in the context of the National Accounts” by: devising output measures that can be reflected in the National Accounts; establishing a clear and transparent relationship between inputs, outputs and thus aggregate government productivity; and establishing the strategic direction for further work in this area, both domestically and internationally¹.

1.2 The Atkinson Review (AR) team were seconded from within government and supported by relevant departments, including the Department for Education and Skills (DfES). As part of the review, we have developed a number of alternative ways of measuring the quality of outputs from the publicly-provided education system. This paper sets out these measures. Whilst we have worked closely with the Atkinson Review team to produce this paper, all views expressed within it are, ultimately, our own.

The paper is structured as follows:

- Atkinson Review and the National Accounts;
- Education output – which education outputs are we trying to measure; what exactly do we mean by education output?
- Current measure of education output and the Atkinson Review recommendations;
- Alternative measures of the quality of education output;
- Comparison of the alternative quality-adjustment measures;
- Valuing education output in monetary terms;
- Extensions to the scope of quality-adjusted education outputs;
- Annex 1: Overview of the National Curriculum System;
- Annex 2: Converting GCSE on to the KS, 2 And 3 APS Scale.

¹ Atkinson Review: Interim Report, Measurement of Government Output and Productivity for the National Accounts (Introduction), July 2004, HMSO.

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

Framework for developing measures of government sector output for use in the National Accounts

1.3 The National Accounts are a record of economic activity in the United Kingdom, including the activities of the government sector. They thus incorporate activities related to the provision of education. Government expenditure and the outputs produced as a result of that expenditure therefore form part of the National Accounts.

1.4 The Accounts capture the volume of output and are intended to indicate broad trends in economic activity over time, so they need to capture changes in government output in a consistent way. Because the Accounts are an indicator, rather than an accurate record of the output from every pound of government expenditure – in our case, education spending – they need to capture the services delivered as comprehensively as possible.

1.5 It is in this context that the final report of the Atkinson Review was published on 31 January 2005, setting out a number of recommendations for improving the measurement of government output, including output from publicly-provided education. It also set out five principles to be applied in the measurement of government outputs:

- Principle A: the measurement of non-market output should, as far as possible, follow a procedure parallel to that adopted for market output.
- Principle B: the output of the government sector should be measured in a way that is adjusted for quality, taking account of the attributable and incremental contribution of the service to outcomes of interest.
- Principle C: account should be taken of the complementarity between public and private output, allowing for the increased real value of public services in an economy with rising real GDP.
- Principle D: formal criteria should be set in place for the extension of direct output measurement to new functions of government.
- Principle E: measures should cover the whole of the UK, and where systems for public service delivery differ and/or data collection differs across different countries of the UK, this variation should be reflected in the choice of indicators.

1.6 Key to the National Accounts is the concept of added-value. We are trying to measure the incremental contribution to output resulting directly from government spending. Moreover, since governments are concerned with achieving ultimate outcomes – as distinct from direct outputs - any measure of what government spending ‘produces’ should ideally capture these.

1.7 For example, the direct output from education spending could be seen as the amount of teaching received by pupils. However, the outcome from that teaching might be the increase in the level of education or knowledge within the population *attributable* to that output – as the final outcome can be

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

affected by things other than government spending on education.

1.8 The National Accounts record changes in the value of output over time. This is straightforward for private sector output, which has a market value – reflected in the prices of that output. But there is no market price for publicly provided education. Nevertheless, the volume of education output should ideally reflect changes in the value of these outputs.

1.9 The Atkinson Review suggested one way in which this could be done, based on the general principle that value of output from government services rises with the real value of private assets and income. Consider the output of education as the acquisition of skills and qualifications. These things have a labour market value, reflected in the earnings associated with different levels and combinations of skills and qualifications. It could then be argued that the value of education output rises with real earnings – which the final report suggests has historically averaged around 1.5% per year in the UK. We also consider another way of capturing the value of education outputs over time.

1.10 Finally, the Atkinson Review recommended that any direct measures of government output need to capture adequately the changing quality of that output. Adjusting for changes in quality has been the cornerstone of our work to date, and is the focus of this paper.

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

EDUCATION OUTPUT

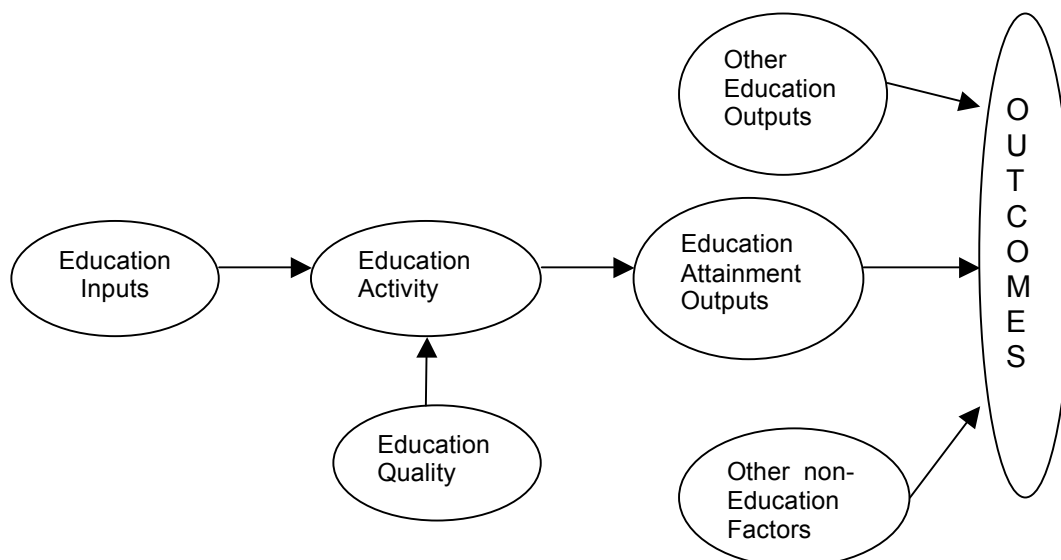
What exactly do we mean by education output?

2.1 Education is a complex process which aims to help achieve a range of ultimate outcomes. In broad terms, these are the maximisation of social and economic welfare: promotion of economic growth; reduction of social inequality; enhancement of happiness or well-being. So when we think about measuring what the education system 'produces', we want ideally to measure the contribution that education makes to these broad aims.

2.2 In the recent Green Paper, *Every Child Matters* (ECM), we identified five specific outcomes that underpin the Department's focus on services to children and young people:

- Being Healthy
- Staying Safe
- Enjoying and Achieving
- Making a Positive Contribution
- Achieving Economic Wellbeing.

2.3 All five outcomes are important and inter-dependent, but can also be affected by things other than government activities. However, measuring the contribution that education makes to each of these outcomes is complex and beyond the scope of our current focus. So the next best alternative is to identify and measure the direct outputs of education that affect these ultimate outcomes. It is these outputs that are the focus of this paper. The relationship between inputs, outputs and outcomes is illustrated below.



2.4 Different education outputs contribute to the five ultimate outcomes. For instance, time spent on sporting activities contributes towards 'being healthy',

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

enjoying and achieving. However, not all of the outputs of education can be measured in the robust and consistent way required for National Accounting purposes – though this does not mean they contribute in any lesser way to the final outcomes.

2.5 For the purposes of National Accounting, we follow international convention on this issue - including Eurostat Handbook guidance – which dictates that education output be defined as teaching activity received, with a suitable adjustment to take account of the quality of that teaching. In many respects, measuring the volume of teaching activity is straightforward. The important, and most difficult task, is identifying ways of measuring changes in the quality of teaching. This problem is the main focus of this paper.

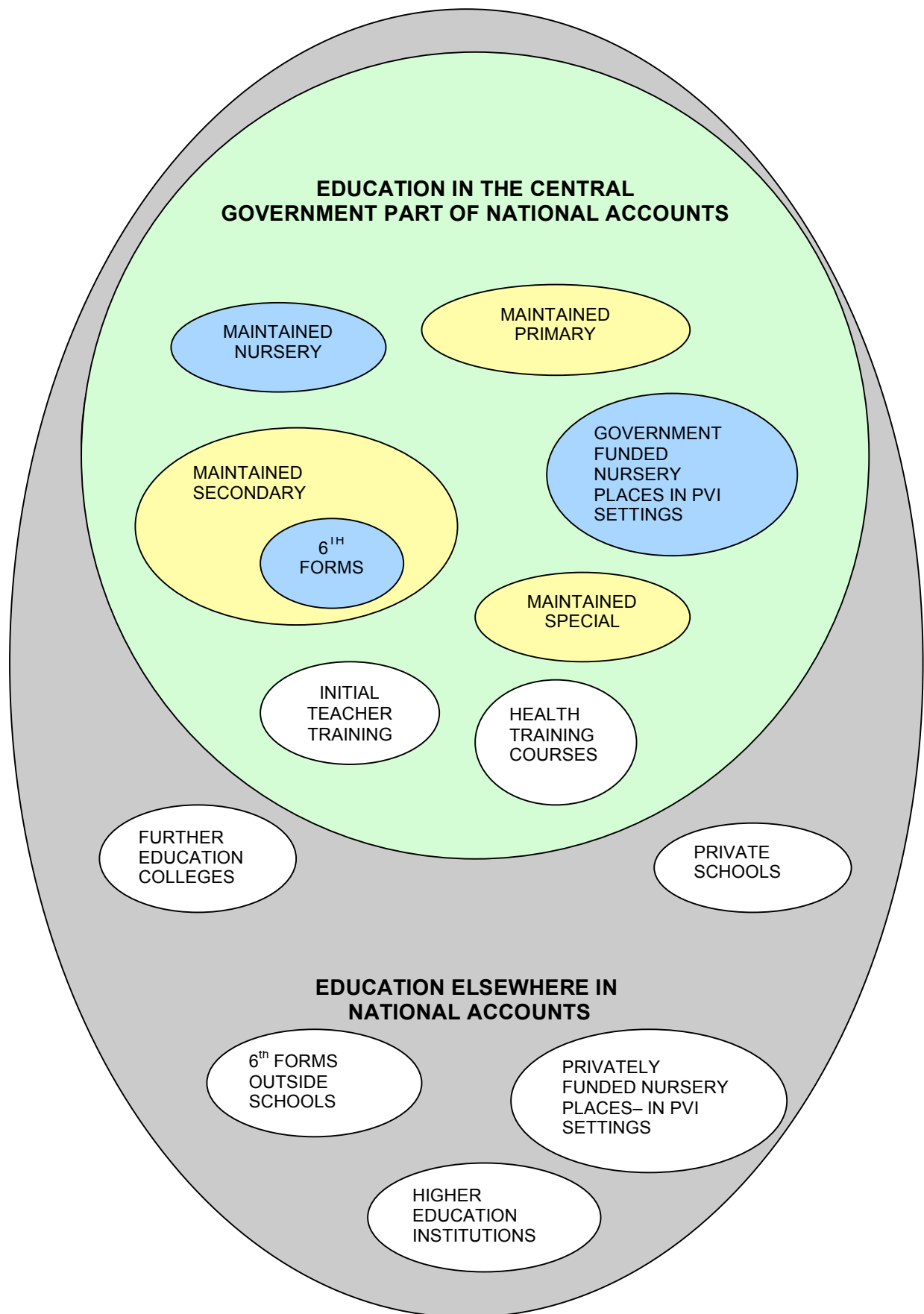
2.6 Our aim is therefore to identify measures of direct, quality-adjusted output which have statistical integrity, and which capture the essence of what education delivers. In the rest of the paper, we discuss the current method of measuring output from the education system used in the National Accounts, and potential ways of improving it.

Which education activities are we trying to measure?

2.7 We have talked about the National Accounts measuring the outputs of publicly provided education, but what does this mean exactly? The figure overleaf illustrates the various education activities provided or funded by the public and private sectors, and where, from 2006, they will be reflected in the National Accounts.

2.8 All the activities surrounded by green are publicly funded and covered in the central government part of National Accounts. These activities were within the scope of the Atkinson Review. Activities highlighted in blue and yellow are those we are concerned with in this paper. Our current focus excludes ITT and Health professional training. We will return to these activities once we have developed robust measures capturing the compulsory phase, with expansion into nursery and sixth forms. The rest are privately funded education activities which are covered in the private sector accounts, and were not within scope of the Atkinson Review. To summarise; the numerical output measures covered in this paper include primary schools, secondary schools and special schools only. We include further discussion of sixth forms and nursery.

MEASURING OUTPUT FROM THE EDUCATION SYSTEM



MEASURING OUTPUT FROM THE EDUCATION SYSTEM

CURRENT MEASURE OF EDUCATION OUTPUT AND THE ATKINSON REVIEW RECOMMENDATIONS

Overview of current measure

3.1 The current method of measuring education output was introduced into the National Accounts in 1998, and is a measure of outputs for the whole of the UK. In accordance with international conventions², output is measured as the volume of education services delivered by nursery, primary, secondary³ and specialist schools, with an adjustment to account for improvements in the quality of education outputs over time.

3.2 To reflect the volume of teaching delivered, the measure uses the number of full time equivalent pupils in each of the four types of school. This volume measure is then adjusted upwards by 0.25% per annum to reflect increasing quality of education output, based on observed past improvements in GCSE exam results. Specifically, it is based on improvements in the Average Points Score of pupils taking GCSEs⁴ in England only, averaged over a period of four years in the mid 1990s. In the current output measure, the rate of quality improvement is therefore assumed to be fixed, with no room to be influenced by policy and changes in spending.

3.3 Since GCSE attainment reflects 11 years of schooling, the eleventh root of the annual percentage improvement is taken. This results from the current assumption that the education process is multiplicative. If an additive model was assumed, then the rate of yearly quality improvements would equal the raw percentage change in the average points score (or roughly, ten times higher than the current 0.25% adjustment).

3.4 This quality adjustment is applied to the volume measure for primary, secondary and sixth forms in secondary schools only. The nursery pupil numbers are not quality adjusted. Each quality-adjusted volume measure is then weighted together by its relative cost and summed. The cost-weighting is based on total UK expenditure for each type of school in the base year of 2000, although the weights were recalculated for Blue Book 2005 using 2002/03 expenditure data.

3.5 This provides a quality-adjusted volume measure of output, which is converted into an index using 2001 as the base year, consistent with Blue Book 2004. The resulting output time-series in the National Accounts for 2004 is illustrated in the table below.

Table 3.1: UK Education Output in the National Accounts ⁵									
	1995	1996	1997	1998	1999	2000	2001	2002	2003
Index	93.9	94.9	96.8	98.0	98.9	99.6	100.0	100.3	100.7

² See Eurostat Guidance and UN System of National Accounts

³ Note that sixth forms in these secondary schools are included in this measure of education output.

⁴ The results are for pupils aged 15 at the beginning of the academic year.

⁵ 'Final Individual Consumption Expenditure of General Government – Education'. Table 6.5 Blue Book 2004. <http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=1143&Pos=2&ColRank=1&Rank=272>.

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

3.6 The current National Accounts measure of education output uses FTE pupil numbers in the whole of the UK, but the quality adjustment applied is based on historical exam results in England only. The Devolved Administrations are examining the potential to apply the methods we describe in their countries, taking into consideration that their education and assessment systems differ by varying degrees from the English system.

Atkinson Review recommendations

3.7 The Atkinson Review made three main recommendations for improving the current method, specifically:

- The **volume** measure should be based on actual school attendance rather than registered pupil numbers;
- The **quality** adjustment should be: revised in terms of methodology and/or updated when new data becomes available; be based on a wider range of information; and should take account of success rates for education provided to pupils aged 16 and over; and
- The measure should account for the changing value of education output in an economy with rising GDP, and we should consider ways of directly **valuing** public sector education output in monetary terms.

3.8 The first recommendation was implemented in Blue Book 2005, and is explained below. The following chapter then explains some alternative ways of measuring improvements in the quality of education outputs. We then go on to discuss ways of measuring the value of education output.

Revised volume measure of output

3.9 The Atkinson Review recommended the use of an output measure based on attendance, as it better reflects pupils who are actually being taught in schools. Two types of data on absence are collected: unauthorised and authorised absences. The aim of adjusting by absence is to get a more accurate measure of the amount of teaching activity received by pupils, so both types of absence are relevant. The following table shows the original volume measure of pupil numbers (based on England only), and shows the same series adjusted by absence. Raw pupil numbers have been declining, but because absence rates have been improving, the overall output index is fairly stable⁶.

Table 3.2 Raw pupil numbers and pupil numbers adjusted by absence							
	1998	1999	2000	2001	2002	2003	2004
Raw pupil numbers ('000 pupils)	6450	6495	6535	6557	6542	6541	6510
Index - raw pupil numbers	98.4	99.1	99.7	100.0	99.8	99.8	99.3
Pupil numbers weighted by absence	5968	6025	6077	6068	6071	6084	6072
Index -adjusted pupil numbers	98.4	99.3	100.2	100.0	100.1	100.3	100.1

⁶ Figures shown from 1998 for consistency with other measures for which data only available from 1998.

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

3.10 The table illustrates the basic volume measure recommended by the Atkinson Review and accepted by the National Accounts Group at ONS. For this analysis, it takes the full-time equivalent number of pupils in England (rather than the UK) who were aged 5 and 15 at the beginning of the academic year, as its base. It covers pupils from Primary, Secondary, Special Schools, City Technology Colleges and Academies⁷. The figures here, and throughout the report, are not directly comparable with those in the Blue Book, and serve to illustrate the methods used.

Issues arising with current quality-adjustment measure

3.11 The table below shows the impact of applying the current 0.25% quality adjustment to the revised volume index.

Table 3.3 Raw pupil numbers and pupil numbers adjusted by absence							
	1998	1999	2000	2001	2002	2003	2004
Index -adjusted pupil numbers	98.4	99.3	100.2	100.0	100.1	100.3	100.1
Quality-adjusted output index	97.7	98.8	99.9	100.0	100.4	100.8	100.9

3.12 Issues arising with the current method of quality-adjustment are:

- a fixed rate of quality adjustment doesn't allow for year-on-year changes in quality. Output therefore only changes significantly if underlying pupil numbers or attendance changes;
- quality is based only on attainment at GCSE level and doesn't reflect changes in attainment at earlier stages in the education process;
- it implicitly assumes that the value of reaching a given qualification level will remain constant over time. Because it is argued that there is a strong causal relationship between education, earnings and productivity, there is an argument that the value of education output increases in line with real earnings;
- all of this means that it is possible to see a fall in overall education productivity (outputs over inputs) simply because education expenditure has risen – as any increase in the real quality of outputs isn't captured, at least immediately. For example, a reduction in class sizes leading to increases in the quality of education would result in an apparent fall in productivity, whilst a doubling of class sizes would suggest virtually a doubling in measured productivity.

3.13 Finally, it raises a linked and difficult intellectual argument about the nature of the education process itself and how we should apply quality adjustments more generally. This issue isn't explained here, but the essence

⁷ City Technology Colleges and Academies are classified as Independent, but for the purposes of this analysis are treated as belonging to the maintained sector.

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

is whether learning is an additive or multiplicative process (or a combination of both). It affects whether the eleventh root (of the % change in APS) accurately reflects the change in the quality of education output.

3.14 The next chapter explains a number of alternative ways of measuring quality over time. Unlike the current measure, we assume that the education process is additive. The Department employs an additive model for its Key Stage test framework. In the absence of evidence to the contrary, we think that this is the more plausible of the two approaches to modelling educational progress. It means that we would not take the 11th root of the raw percentage changes in the alternative quality measures – in APS, for example – but the raw changes in education quality implied by the different approaches.

Revising the output index to reflect the value of education

3.15 The Atkinson Review Final report points to the complementarity between public and private services – by which we mean that the value of public sector services may change as the real value of people's income and assets change. The AR recommends that public sector output measures should take account of this complementarity, and should reflect the increased real value of public services in an economy with rising GDP (Principle C).

3.16 In the case of education, output is proxied by the level of knowledge pupils attain, as measured through qualifications. The value of these qualifications rises in a climate of real wage growth. The AR recommends that we give serious consideration to adjusting the education output measure to reflect rising earnings. More specifically, it suggests that, historically, real earnings in the UK have grown by around 1.5% per year. Therefore, an adjustment of 1-1.5% per year to education output would seem appropriate.

3.17 Table 3.4 shows that this adjustment increases the gradient of the output index. Chapter 5 goes on to explain an alternative way of valuing education output, based on the additional lifetime earnings gain from having the qualifications that children can take at the end of compulsory education.

Table 3.4 Revised volume measure adjusted for real earnings growth							
	1998	1999	2000	2001	2002	2003	2004
Revised quality-adjusted volume index	97.7	98.8	99.9	100.0	100.4	100.8	100.9
With 1.5% real earnings growth pa	93.3	95.9	98.5	100.0	101.9	103.8	105.5

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

ALTERNATIVE MEASURES OF THE QUALITY OF EDUCATION OUTPUT

4.1 Ideally, we should measure the quality of education outputs from each part of the education system, but data limitations mean this isn't possible. So we need a measure that suitably proxies quality of the education system as a whole.

4.2 We consider three ways of measuring quality:

- GCSE attainment at the end of compulsory schooling;
- Pupils' progress which pupils between each of the four Key Stages of the English compulsory education system;
- Results of Ofsted inspections.

4.3 All three methods better reflect year-to-year changes in the quality of education 'produced', in contrast to the fixed quality adjustment factor currently used in the National Accounts. We illustrate each quality measure in its index form, and also multiplied by the Atkinson-revised volume index to derive a quality-adjusted output index.

GCSE ATTAINMENT OUTPUT METHODS

4.4 Pupil attainment is perhaps the most obvious and transparent measure of education output, as it measures pupil knowledge and it is measured at different stages in the school system. We focus first on performance at the end of compulsory education – i.e. on GCSE attainment.

4.5 There are two potential ways of measuring quality using GCSE results:

- proportion of GCSE pupils who achieve a threshold level of attainment, taken to be 5 or more A*-C grades; or
- average GCSE point scores (APS).

Threshold attainment method

4.6 This method is very simple. We take the annual proportions of 15-year-olds in each year who achieve 5 or more A* to C grades in their GCSE exams (or equivalents), and convert them into an index. The impact of the quality index can be seen by comparing the Atkinson-revised volume index (pre- quality adjustment) with the output index (post- quality adjustment).

4.7 Unlike the current approach – which applies a fixed adjustment based on the average change in GCSE attainment – our measure is based on the annual changes in the percentage gaining 5 good GCSEs. So the quality adjustment for a particular year is based on the percentage gaining 5 good GCSEs in that year.

Table 4.1 GCSE Threshold Measure							
	1998	1999	2000	2001	2002	2003	2004
Atkinson-revised Volume Index	98.4	99.3	100.2	100.0	100.1	100.3	100.1
Quality Index	91.4	95.4	98.3	100.0	103.3	106.7	108.6
Output Index	89.9	94.7	98.5	100.0	103.4	107.0	108.6
With 1.5% real earnings growth	86.0	92.0	97.0	100.0	104.9	110.2	113.6

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

Average Points Score Method

4.8 This quality-adjustment index is similar the threshold measure above, but is derived from a time-series of the average number of points that students get in each of their subjects at GCSE level. GCSE points are summed across all individuals, for all the subjects that they were enrolled for at the beginning of their GCSE exam year. The measure is then based on taking the average of these points scores across all pupils in that cohort.

Table 4.2 GCSE APS Measure							
	1998	1999	2000	2001	2002	2003	2004
Revised volume Index	98.4	99.3	100.2	100.0	100.1	100.3	100.1
Quality Index	91.6	94.8	97.9	100.0	101.0	103.4	105.5
Output Index	90.1	94.1	98.0	100.0	101.1	103.7	105.6
With 1.5% REG	86.2	91.3	96.6	100.0	102.6	106.8	110.4

Issues arising from the GCSE attainment methods

4.9 Both of these GCSE attainment methods improve on the current method, by reflecting changes in the quality of education outputs over time, rather than averaging over a number of years to obtain a fixed rate of quality improvement. However, they also raise a number of issues, largely relating to how well even these improved measures capture the quality of outputs from the education system in any given year:

- The GCSE threshold may not reflect changes in learning over the full range of ability levels. Improvements in the attainment of low-achieving pupils will not be reflected if these pupils lie far below the 5A*-C threshold: they will move towards the threshold but may not cross it. Similarly, improvements for high-achieving pupils will not be picked up if they would have achieved 5A*-C anyway;
- A potential difficulty with any quality measure based on attainment is that the scales used for measuring attainment might not exactly measure changes in pupil learning. A 2% increase in GCSE APS between years might not necessarily mean that pupil learning (and, with it, the quality of education) has increased by exactly 2% - unless it has an exactly proportional relationship with the APS scale;
- A related issue is the non-linearity of the APS system itself. Under the old GCSE APS system a grade G is counted as 1 point and a grade A* is counted as 8 points. It doesn't necessarily mean that someone with an A knows 8 times as much as someone with a G. This will affect percentage changes;
- Our scales for measuring attainment have changed over time. The implication is that different quality indices could be produced for the

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

same years of exam results, depending on the scale used to measure attainment. This could apply even where the qualifications counted as attainment remain unchanged;

- A fundamental concern is that changes have, and may continue to be made, in what is counted as exam attainment. For example, a new measure of GCSE/GVNQ attainment was introduced for 2004 including a wider range of other approved pre-16 qualifications and the new GCSEs in vocational subjects⁸. Changes in the scope of what is counted could show up as increases in output – but the impact is small (but still significant) at around 0.1% in 2004. The implication is that time series may suffer from inconsistency over time. Further work is needed to assess the implications of these changes fully;
- A potential difficulty for both approaches – and the current method - is that they use exam attainment of only one cohort of pupils in education in any given year (i.e. those sitting GCSEs) to proxy quality changes across the whole pupil body in that year. Moreover, GCSE attainment for this cohort will implicitly reflect a pupil's education over all 11 years of schooling (on average), and not just the current year of their GCSE exams. The implication is that GCSE quality measures – used as a proxy – are still insensitive to concurrent changes in quality throughout the rest of the education system;
- Threshold measure – and to a less obvious extent, the APS measure – is subject to ceiling effects, for example, as the % achieving 5 good GCSEs approaches 100%. Since the present level is around 50%, it will be many years before the ceiling effects matter in aggregate;
- However, there will be a compositional effect, as the ceiling will be felt much sooner at school level – some schools already have a high percentage of pupils above the threshold and so have much less scope for improving performance (increasing educational output). This is more of an issue for disaggregated analysis by school or local area;
- The APS measure could be affected by any changes in school policy on admissions for examinations – for example, entering pupils for increasing numbers of GCSEs. Though arguably, if qualifications are a key education output, changes in the number of GCSEs taken should rightly be included in any output measure;
- Finally, an attainment quality measure based on exam results will be affected by changes in the examination framework. Such changes in

⁸ Though time series for this article are based on the old metric, provisional results suggest that the inclusion of wider qualifications has increased the national APS by approximately 3 points in 2003/4⁸ – equivalent to half a grade at GCSE. As the scope of qualifications included widens, the gap is likely to increase. These issues have implications for the consistency of any time-series data. Provisional results also suggest that the GCSE threshold measure has changed by 0.1 percentage points with the introduction of more qualifications. For schools and LEAs where the take up of specific "other" qualifications is higher, the differences are more notable. In future years as the take up of the wider range of qualifications increases we would expect this difference at national level to increase.

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

the GCSE data collection occurred in 2004.

4.10 There are ways of addressing some of the issues raised above:

- It would, in principle, be possible to develop a GCSE threshold measure based on the proportion of pupils crossing several thresholds (e.g. passes only at D-G, 1-4 A*-C, 5 A*-C etc.). Such an approach – not developed or presented here – would also need to consider how to weight together the different thresholds of GCSE categories to produce a single index. One way of doing this might be to base the weights on the valuation of those attainment levels in the labour market, if the necessary data were available;
- That attainment and pupil learning might not be proportional is not unique to the GCSE APS method. The key issue is therefore to assess how severe its implications are for the accuracy of quality measurement, and how, if at all, it bears on our choice of measure;
- Ideally, we would want a measure that captured the progress made by all pupils in education in any given year, and only that progress attributable to the education they receive in that year. The total progress in each year would be summed over all pupils, reflecting progress made by all cohorts and throughout the education system. It would address the problems outlined above but data does not exist to support such a measure – we would need test results for each cohort of children at the end of each year.

4.11 Despite the limitations, these GCSE methods can be measured annually and are relatively simple. Employers value qualifications, and on this basis the GCSE-based measures are appropriate output measures. The threshold method also has potential scope for allowing education outputs to be valued in monetary terms – something discussed in detail in Section 6. The next section explains an alternative approach aimed at addressing some of key drawbacks outlined here.

COHORT PROGRESS OUTPUT METHOD

4.12 Compulsory education in England is divided into four Key Stages, and pupils sit national examinations at the end of each phase. Further information about the English National Curriculum System is given in Annex 1. Data on these examinations allows us to derive a more refined measure of the quality of education outputs based on more than one phase in the education process.

4.13 A measure of average cohort progress over an entire Key Stage can be derived as the difference between (i) the average points score for a given cohort of pupils at one Key Stage and (ii) their average points score at the next Key Stage. A further refinement to this method – not developed here – is to use the National Pupil Database to link the attainment of individual pupils across the four Key Stages. From this we could derive a measure based on the actual progress made by every pupil – i.e. essentially the sum of individual

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

progress made between stages, rather than the average.

4.14 The Key Stages vary in length. The longest Key Stage is four years – Key Stage 2. The earliest year for which cohort progress measures are available is 1998, as the Key Stage APS data are only available from 1995 onwards, the year in which Key Stage testing was introduced.

Measuring cohort progress

4.15 Cohort progress is calculated as the simple arithmetic difference between the average points score (APS) at each Key Stage and the APS at the preceding Key Stage *for the same cohort of pupils*. These data are shown in the table below and reflect the progress, over the whole of a particular Key Stage, for the cohort sitting the exams in that year.

4.16 Key Stage 2 is the longest of the four Key Stages and lasts four years. In 1999, cohort progress at KS2 was 11.2 points. This measure of progress – or added value - is based on the test results for only those pupils who took their KS2 tests in that year alone. It does not capture progress made by the other three year groups who were in KS2 in 1999. At the simplest level, these measures of cohort progress are weighted together - to form an index of quality - according to pupil numbers in each of the Key Stages (table 4.4).

Table 4.3 Simple Cohort Progress (Key Stage points)				
Year	KS1 progress	KS2 progress	KS3 progress	GCSE progress
1998	5.6	10.6	7.8	5.3
1999	5.9	11.2	7.5	5.8
2000	6.2	12.6	7.3	6.0
2001	6.4	12.5	7.9	6.3
2002	6.5	12.7	7.2	6.1
2003	6.5	12.4	7.1	5.8
2004	6.4	12.3	6.9	5.9

Table 4.4 Simple Cohort Progress Measures							
	1998	1999	2000	2001	2002	2003	2004
Volume Index	98.4	99.3	100.2	100.0	100.1	100.3	100.1
Quality Index	88.1	91.5	96.6	100.0	98.3	96.0	95.0
Output Index	86.7	90.9	96.8	100.0	98.4	96.3	95.1
With 1.5% REG	82.9	88.2	95.3	100.0	99.8	99.2	99.4

Issues arising with a simple cohort progress measure (CPM)

4.17 The main advantage of this simple cohort progress measure is that it is based on the attainment outcomes of 4 cohorts of pupils, rather than one cohort, as in the GCSE methods. It is therefore closer to a measure of the quality improvement of the education system as a whole.

4.18 There are, however, a number of disadvantages:

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

- Progress made by the other seven year groups is not covered. However, we have derived a method for adjusting the data to include estimates of their progress in a given year – described below;
- In the absence of data, we assume that pupils enter compulsory education with 9 National Curriculum points⁹. This is the comparator against which KS1 progress is measured – but it assumes that prior attainment is fixed. It means, for example, that increases in prior attainment due to improved Early Years provision will be credited incorrectly to changes in the compulsory school system. This gives rise to a potential inaccuracy in the CPM;
- In addition, the assumption of 9 National Curriculum points could be restrictive even where prior attainment is fixed between years. If, for example, pupils enter the system with 10 National Curriculum points, then the level of progress we estimate for KS1 using a prior attainment of 9 points will be too large. As a result we could underestimate the true percentage changes in KS1 progress, given that the denominator used for those percentages would be too large. The scope for determining a more exact estimate of achievement pre-school is discussed in section 9;
- A further, key, issue with the Cohort Progress Measure is that GCSE scores for years 1994 to 2004 were measured on a different points scale from that used to capture attainment at each of the earlier Key Stages 1 – 3. To allow a measure of cohort progress from KS3 to KS4, we thus have mapped GCSE points onto the KS3 APS scale, by a process described in Annex 2. However this mapping involves a statistical restriction which, like the prior attainment assumption discussed above, could affect the percentage changes calculated for progress at KS4. This restriction is treated in more depth in the Annex, although we suspect its impact will be slight;
- Using cohort means could affect the interpretation of the CPM, as it ignores losses to or gains from the private education sector (or from immigration and death). This may be a particular issue when looking at progress between KS2 and KS3, as this is the point at which pupils switch to private education at secondary level. For example, if the brighter pupils leave the state system after KS2, their good performance prior to that will be captured in the KS2 results, but they will be absent from the KS3 results. This could bias downwards the measured progress between KS2 and KS3;
- This issue may provide grounds for using matched individual pupil level data to measure progress between Key Stages – and it would be

⁹ The Curriculum System assumes that an average pupil will progress at a rate of one point per term. As the expected level of attainment for KS1 is 15 points, and KS1 is two years in length, or worth 6 points, we obtain a level of 9 points on entry by subtraction.

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

useful to explore how sensitive the measure is to using means or individual level data;

- Comparing differences in pupil progress across years could be affected by changing variabilities in KS scores, especially if there are ceiling effects. One option would be to standardise results to fix the variability over time, although this would result in a different conclusion about relative progress than using raw scores;
- Measured progress between KS3 and KS4 is dependent on the mapping we use to transform GCSE scores onto the KS scoring scale. More exploratory work is needed to test how sensitive the results are to the transformation model used, in terms of sample selection (which year used to estimate relationship; parameter constancy) and functional form (linearity);
- Need to ensure consistency of the quality measure based on CPM with the underlying volume measure of output, which is based on absence-adjusted pupil numbers.

Adjusted cohort progress

4.19 One way of moving further towards a measure reflecting all 11 pupil cohorts in any given year, is to interpolate the progress made by the cohorts not sitting exams from the results of those cohorts who did sit exams. To do this we have to make assumptions about how much progress is expected to be made in each year of a Key Stage. The simplest assumption – and the one used here – is that pupils make even progress throughout each stage.

4.20 Consider an example. KS2 progress in 2001 was 12.5 points. Since KS2 lasts for 4 years, the average progress made in each year of the stage is assumed to be around 3.1 points - for that cohort. So we assume that this cohort made progress of 3.1 points in each year of their KS2 – from 1998 to 2001. We repeat this process for all KS results, giving a matrix of un-weighted estimates of the *annual* progress made in each year, for each of the 11 cohorts. The table below illustrates this for KS2, with the diagonals representing the progress of each cohort.

Table 4.5 Key Stage 2 - Adjusted Cohort Progress					
		1998	1999	2000	2001
KS2	3	3.1	3.2	3.1	3.1
KS2	4	3.1	3.1	3.2	3.1
KS2	5	2.8	3.1	3.1	3.2
KS2	6	2.6	2.8	3.1	3.1

4.21 For each academic year, we then use pupil numbers to weight together the estimates of progress for each cohort (i.e. weighting the rows in each column). This gives more or less weight to a particular cohort depending on

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

its size. The resulting measures of *annual* average progress are illustrated below Key Stage 2, for the years 1998 to 2001. For comparative purposes, we also show the implied, artificially constructed estimates of total Key Stage progress (assuming equal progress in each of the 4 years). The same process would be repeated for each of the Key Stages to derive a measure of output from primary and secondary education. The results are translated into a quality index using 2001 as the base year.

Table 4.6 Cohort Progress for Key Stage 2				
	1998	1999	2000	2001
Average yearly progress	2.9	3.1	3.1	3.1
Total progress (4 years)	11.7	12.2	12.6	12.5

4.22 This seems like a complicated approach, but it demonstrates the logic behind a method based on capturing progress (added value) throughout the education system in any given year. We can demonstrate that essentially the same results can be derived by simple averaging of the total KS results¹⁰. The example below illustrates this for KS2.

Key Stage 2	
Year 3 – measured in 2002	12.7
Year 4 – measured in 2001	12.5
Year 5 – measured in 2000	12.6
Year 6 – measured in 1999	11.2
Average	12.2

Estimates of cohort progress beyond 2001

4.23 One drawback of this extension of the cohort progress method is that the adjustment process becomes less complete for the most recent years of a time series. For example, we won't know until 2005 the actual KS2 attainment of those in their first year of KS2 in 2002 – so we have no a priori basis for distributing their annual attainment backwards. By 2004, the problem becomes particularly serious as we only have data for four out of the eleven cohorts - it is only 36% complete.

Table 4.7 The number and percent of year groups for which we have data		
	Number of year groups for which we have results using adjustment	Percentage of year groups for which we have results using adjustment
1998	11	100%
1999	11	100%
2000	11	100%
2001	11	100%
2002	10	91%
2003	8	73%

¹⁰ Weights used to account for cohort size are fairly stable over time, so have little impact on the results.

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

2004	4	36%
------	---	-----

4.24 For National Accounting purposes we therefore need to make estimates of annual cohort progress beyond 2001, for pupils sitting their KS exams in later years. So the quality (and output) index would be subject to revision by up to 4 years into the future (as KS2 is the longest stage). The shaded cells in the table below illustrate the data points that we need to estimate for Key Stages 1 and 2. The estimates are based on taking the average of the available data points for that Key Stage, in that year.

Table 4.8 KS2 - Adjusted Cohort Progress								
		1998	1999	2000	2001	2002	2003	2004
KS2	3	3.1	3.2	3.1	3.1	3.1	3.1	3.1
KS2	4	3.1	3.1	3.2	3.1	3.1	3.1	3.1
KS2	5	2.8	3.1	3.1	3.2	3.1	3.1	3.1
KS2	6	2.6	2.8	3.1	3.1	3.2	3.1	3.1

4.25 The table below shows the resulting measures of annual cohort progress, and their translation into a quality index. The estimates of average yearly progress in primary and secondary are weighted by the volume of pupils in each phase of education¹¹.

Table 4.9 Cohort Progress Measure							
	1998	1999	2000	2001	2002*	2003*	2004*
Volume Index	101.2	101.5	101.2	100.0	98.9	97.9	96.8
Average yearly progress	2.78	2.89	2.95	2.96	2.91	2.88	2.86
Quality index (weighted)	93.9	97.6	99.7	100.0	98.3	97.3	96.6
Output index	92.4	96.9	99.9	100.0	98.4	97.6	96.7
With 1.5% REG	88.4	94.1	98.4	100.0	99.9	100.5	101.1

* Provisional estimates

OFSTED ASSESSMENT METHOD

4.26 It is not necessary for a quality-sensitive measure of education output to be based on attainment. In principle, Ofsted school inspection results could be used to proxy changes in the quality of the education system. We consider the advantages and disadvantages of using a judgement on 'overall school effectiveness', and also of using a composite measure combining a number of individual assessment categories.

Overall School Effectiveness Measure

4.27 Since 2000, Ofsted inspectors have been asked to judge the 'overall effectiveness' of a school. This takes into account the school's strengths and weaknesses, the context of the school, and how it performs compared with similar schools. It is based on a careful review and assessment of all the evidence gathered about a school.¹² This judgement is graded on a scale of 1

¹¹ The weighting here is based simply on the number of years in the primary stage (6 years) and secondary stage (5 years) as a proxy for the difference in pupil numbers in the two stages.

¹² <http://www.ofsted.gov.uk/publications/docs/hb2003/primaryhb03/hmi1359-09.html>

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

to 7.

4.28 To derive the quality index, the school results of 'overall effectiveness' have been aggregated to national level. The total number of primary schools achieving each grade is summed and presented as a percentage of all primary schools. That percentage is then weighted so that good results receive a higher score – reflecting higher quality - than lower results. Essentially we are deriving a weighted average score for primary school effectiveness, based on the proportion of schools in each grade. For the purposes of this analysis the weighting was done by reversing the grading scale, so that the highest scores are weighted by 7 and the lowest by 1. Each weighted percentage is summed to give an overall score for all primary schools. The same methodology is applied to secondary schools.

4.29 The primary and secondary scores are then weighted by the number of pupils in each type of school, as a percentage of all schools. The weighted scores are summed to give a result for all schools. The resulting quality index is applied to the volume measure of pupil numbers adjusted by absence data. Table 4.10 shows the resulting output index.

Table 4.10 Ofsted 'Overall Effectiveness' measure					
	2000	2001	2002	2003	2004
Volume Index	100.2	100.0	100.1	100.3	100.1
Quality Index	98.1	100.0	101.2	101.5	99.1*
Output Index	98.3	100.0	101.2	101.8	99.2*
With 1.5% REG	96.8	100.0	102.7	104.9	103.7

* The framework of assessment changed between 2003 and 2004 making the time-series inconsistent. See paragraph 7.12

Issues arising from the overall school effectiveness measure

4.30 A measure based on Ofsted results has the advantage of encompassing all judgements that inspectors make about a school. It has two implications. Firstly, the coverage of school outputs is wide – it will include not simply attainment, but teaching and behaviour etc. Secondly, as the inspector's judgement combines all elements of a school in one metric, it is not necessary for us to weight together each dimension of school performance.

4.31 However, it also has a number of disadvantages:

- A weighting has to be applied to the grading scale. Although the scale is numerical (from 1 to 7) it is not necessarily linear. So a school given a 2 is not necessarily twice as 'effective' as a school given a 4;
- The framework of assessment has changed a number of times since Ofsted's inception. The overall effectiveness judgement has only been used since 2000, was changed in September 2003, and again in September 2005, to a 4 point scale. This highlights the problems

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

associated with deriving time-series data;

- Any Ofsted measure risks being sensitive to the weighting used by Ofsted to gross up individual school results to a national level. Ofsted Inspections are not a random sample so raw inspection results each year are skewed. When Ofsted present national figures they take into account the types of schools inspected that year, and aim to adjust to present a nationally representative picture;
- Generalised concern that Ofsted assessments are not valid for what we want to capture in the National Accounts – which is the quality of education outputs, i.e. qualifications;
- The time series is not very long, and changes in the assessment framework in 2003 and 2005 are likely to have caused the drop in the index between those years, demonstrating discontinuity;
- In addition, the Atkinson Review Interim Report notes that;
 - Schools prepare thoroughly for Ofsted inspections and there may be an artificial climate, and no certainty that the criteria for teaching and school quality would be consistent across schools and time. (Although the reduction in the period of notice since September 2005 has to a large extent addressed this issue)
 - Ofsted's remit is not the same as needed for National Accounts, i.e. a measure of value-added by the school for individual pupils.

Combined Ofsted Measure

4.35 Ofsted also make judgements about individual elements of a school, namely, so in principle, a composite measure of the quality of a school could be derived from combining results in the following categories:

- Quality of teaching;
- Pupil's achievement;
- Leadership and management;
- Personal development (including attitudes and behaviour).

4.35 This approach, however, poses a number of potential problems:

- Most importantly it would require an arbitrary weighting between the different components. For instance, should teaching be judged to be more important, and thus have a greater weighting than management?
- A further weakness is the disparity between a composite measure and the 'overall effectiveness' judgement. Although both are collected by the same inspection team they can produce inconsistent results. Inspectors make a judgement about the schools 'overall effectiveness' based on evidence which may not be wholly represented in the

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

individual assessment categories

- Due to these limitations and the generic weaknesses outlined above, we have not developed this measure further.

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

IMPACT OF ACCOUNTING FOR QUALITY

5.1 This section briefly summarises the impact that each of the quality measures – including the current approach - has on the final index of education output. We ought to be able to understand any patterns in the output indices, not least as a way of providing some kind of intuitive validity to the approaches. Note again that all the indices presented in this paper are not an exact depiction of what would appear in the National Accounts, but they do accurately reflect what the indices would look like for English primary and secondary school education. All quality-adjusted output indices – including the current method – include a 1.5% adjustment to reflect increases in the real value of education outputs.

Table 5.1 Comparison of the different quality adjustment methods							
	1998	1999	2000	2001	2002	2003	2004
Volume Index	98.4	99.3	100.2	100.0	100.1	100.3	100.1
Quality-adjusted output indices							
Current method	93.3	95.9	98.5	100.0	101.9	103.8	105.5
GCSE threshold method	86.0	92.0	97.0	100.0	104.9	110.2	113.6
GCSE APS method	86.2	91.3	96.6	100.0	102.6	106.8	110.4
Simple cohort progress method	82.9	88.2	95.3	100.0	99.8	99.2	99.4
Adjusted cohort progress method	88.4	94.1	98.4	100.0	99.9	100.5	101.1
Ofsted inspection method	n/a	n/a	96.8	100.0	102.7	104.9	103.7

* Revised volume measure is not yet reflected in the National Accounts

5.2 The key points to note are:

- The basic volume index is quite flat, particularly after 2000. Although pupil numbers have been falling, attendance rates have risen, and these two effects have balanced each other out;
- Accounting for the quality of education outputs has quite a significant effect on the output index. This means that it is vitally important for any quality measure we use in the National Accounts to be a fair reflection of the true quality of education output;
- All of the quality-adjustment methods lead to an increasing output index up to 2001, but have differential impacts beyond that point;
- The Ofsted-based index is fairly flat, with some small fluctuations. Overall, it is not very significantly different to the basic volume index;
- The three GCSE-based indices – including the current method – all show a continuing increase from 1998, although the effect is much more muted for the current method; and
- The cohort progress based indices have a very different and distinct pattern – rising before falling after 2001.

VALUING EDUCATION OUTPUT IN MONETARY TERMS

Current valuation of education output in the National Accounts

6.1 We have been discussing education output in terms of a volume index (or series). A volume series is most simply expressed in index number form, referring to a base year which takes the value of 100. If growth between the base year and the following year was 0.3%, the index would show a move from 100 to 100.3. National Accounts convention also expresses this type of index in money terms. In such a case, the value of the series in the base year – or reference year, to use a more precise term - is the same for both the volume index and the money flow which underlies it.

6.2 To take the education example, Blue Book 2004 (BB04) shows an amount of £33,480 million for general government final consumption expenditure on education at current prices in 2001. This happens to be the reference year. The volume measure shows the same amount for 2001. Growth between 2001 and 2002 was 0.319%; hence the volume series for 2002 shows a value of $1 \times 1.00319 = £33,587$, based on Reference year 2001.

An alternative way of valuing education output

6.3 Chapter 3 identified one way of taking into account changes in the value of education output. Public sector education output could be valued more accurately to reflect the value that society and the economy places on it - as is done for private sector output. Education – according to one theory – increases the productivity of the workforce, which is reflected in increased earnings and leads to rising living standards.

6.4 One intuitive way of valuing education output is therefore through the labour market value of qualifications, over the lifetime of individuals possessing them. In theory, this can be measured as the estimated lifetime earnings of individuals with different levels and combinations of qualifications. Strictly speaking, the National Accounts are concerned with the added-value of economic activities, so we need to estimate the added-value of education.

6.5 The added-value of education output can be seen as the difference in lifetime earnings of those who have had no education, with those educated or qualified to a particular level. So for example, if someone left compulsory education with five A* - C at GCSE level, we could derive a measure of value-added by comparing the estimated lifetime earnings of those with five GCSEs at A*-C with the earnings of people with no education¹³.

Estimating lifetime earnings

6.6 Data limitations mean that we can't easily put a labour market value on

¹³ We ideally want to make comparisons with their counterfactual earnings, but we can't observe these.

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

all combinations of qualifications on leaving school. Equally important, because we do have a compulsory education system, there are very few individuals who have never been to school. So we can't observe the earnings of a completely uneducated person. The next best proxy – though not without intellectual problems – is the earnings of a person with no qualifications.

6.7 A recent academic study found that – for example – the wages of those with O-Levels were around 18% higher than those with no qualifications¹⁴. Using this figure – together with the earnings of the no qualification group – we can estimate the average additional lifetime earnings associated with having O-levels.

6.8 We also need to take into account various other factors, such as: probability of being in employment; real earnings growth; and non-wage labour costs (which also reflect the labour market value of employees). The resulting additional earnings profile is *discounted* to reflect that the additional earnings accrue in the future. So the value in the National Accounts would reflect the present-day value of education output. The following table summarises these adjustments.

<u>Adjustment</u>	<u>Argument</u>	<u>Value</u>
Real Earnings Growth	Real earnings tend to grow over time	1.5%
Discounting	The lifetime earnings profile needs to be valued in present value terms. 3.5% is the HMT's current discount rate	3.5%
Employment probability	Earnings profiles are deflated by the probability of employment at each qualification level, to account for the fact that workers will not all be employed for their full working life. Employment rates vary according to level of highest qualification held	Various
Non-wage labour costs	A rational employer would not pay an employee more than the value of their productive output. This includes non-wage labour costs	25%

6.9 Using this methodology, a conservative estimate of the additional lifetime earnings from having five A*-C GCSEs (using 2003 data) is around £275,000 – for one individual. We could aggregate this up to the number of pupils leaving school with five GCSEs at A*-C to derive an estimate of the total value of education (as reflected in GCSE results). It is clear to see that this method might be more easily applicable to the GCSE threshold measure. National Accounts need to reflect changes over time. So any approach would need to accurately reflect annual changes in the value of education – something which is in practice likely to be difficult.

Data and methodological issues

¹⁴ The academic study is by Blundell, Dearden and Sianesi (2003) using the National Child Development Survey. 18% is the average difference in wage rate over the life-cycle.

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

6.10 There are two opposing influences at work that could bias this estimate of the value of output:

- a) Currently more than half of pupils achieve 5 or more A*-Cs at GCSE. Since this is the highest GCSE qualification level measured in the LFS, any improvement in standards within this category (e.g. resulting from policy targeted on gifted and talented children) will not show up as an increase in attainment. This would lead to *undervaluation of increases in the output* of the education system. The ONS are currently piloting an additional LFS category of 8+ A*-C, which would cover the top third of pupils at present;
- b) individuals who cross the threshold to a higher category of attainment would be attributed an income premium equal to the full difference in average income between the attainment groups, which could be an *over-estimate* of the true added value of education;

6.11 The two above effects offset each other, but it does not necessarily follow that the two effects cancel each other out completely. This estimate underplays the true benefits of education for a number of reasons:

- the wage premium used is based on a comparison of wages for those with O levels and those who leave school without qualifications. But this may underestimate the true return to education, since pupils leaving school without qualifications may have benefited from many years of education and, arguably, would earn much less if they had never been to school. This suggests a much larger value of education;
- wage benefits for pupils achieving less than 5 A* - C at GCSE have not been counted, yet these individuals may enjoy substantially higher earnings than they would have done without education;
- education confers many other benefits which are necessarily excluded from our calculations. For example, evidence suggests that educated people live a healthier lifestyle, costing less to the NHS and enjoying a better quality of life. In addition they are less likely to be involved in crime, reducing the burden to the criminal justice system and avoiding unnecessary hardship to the victims of crime;
- education is an enjoyable experience for many pupils contributing to all aspects of their life, while providing a childcare service to parents which allows them to return to the work place. Preliminary calculations put the value of this childcare service in the region of £15bn to £20bn per annum, based on the value parents attach to childcare in terms what they are willing to pay for it on the open market.

6.12 Such labour market valuation is useful in that it is an absolute and direct measure of the value of education output. It implicitly incorporates the quality of those outputs, because if this changed, so would its labour market value. However, a key limitation of the method is the time lag needed to properly detect changes in the *lifetime* value of education outputs, as we have

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

to estimate this over a 40 year working life.

EXTENDING THE SCOPE OF THE OUTPUT MEASURES

7.1 So far we have examined the measurement of quality-adjusted outputs from compulsory schooling, but the focus of the National Accounts is on *all* state-provided education. The output of sixth forms is currently measured by the volume of pupil numbers in sixth forms, adjusted for quality by 0.25% in the same way as the volume measures of primary and secondary schools. In contrast, the output of nurseries is currently measured by pupil numbers, but there is no adjustment for quality. In this section we explore ways of measuring quality in nurseries and sixth forms, but at this stage we have not attempted to translate them into numerical output indicators.

Early Years education

7.2 There are two main types of state funded nursery education; that which is provided by the state – maintained nursery schools and classes; and that provided by the private, voluntary and independent (PVI) sector, but which is partially funded by the state through subsidised places for 3 and 4 year olds. The latter has traditionally been viewed as outside the scope of government expenditure in National Accounts, and output from the former has been included. The Atkinson Review recommends the inclusion of both. In the interim, the Final Report further recommends that their output should be included on the basis of pupil numbers, and although a quality adjustment will not be applied immediately, it should be considered for future inclusion.

7.3 There are a number of options for deriving a quality measure of nursery education. We first consider how the quality measures derived for compulsory schooling could be applied to early years. We then consider specific quality measures derived from early years education itself.

7.4 The simple application to early years of a quality measure derived from attainment during compulsory education raises a number of issues. Under both GCSE methods, the quality of early years is assumed equal to that of KS4, which is obviously restrictive. Moreover, a GCSE-based quality method would not register any improvements in early years quality the cohort(s) affected reach KS4, approximately ten years later. Even then, any impact of early years quality changes would have to persist to be picked up at all.

7.5 The pupil progress methods look at attainment across key stages, and in theory it would be possible to extend these methods to early years education – using the Foundation Stage Profile (FSP). We could use FSPs to measure progress between the beginning and end of the foundation stage. Foundation stage starts at age 3, in nursery, to age 5 in the reception year in primary school¹⁵.

7.6 The assessment of attainment of children prior to KS1 is difficult, not least because of the young age of the children. However, foundation stage

¹⁵ Schooling becomes compulsory in the term after a child turns 5.

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

profiles – where a professional judgement is made by teachers about a child's achievements and their progress - were introduced in 2003. The foundation stage comprises of thirteen assessment scales in six areas, each of which has nine points.

7.7 The FSP have three disadvantages:

- Importantly, FSP are not consolidated into one scale, and are not given on a comparable scale to key stage results. So it is not possible to link directly achievements at FSP with those at KS1, and therefore derive progress, as part of the pupil progress method;
- FSP are not collected at a national level and only 10% of results are centrally collected¹⁶. Results would require grossing up from sample data;
- Finally, many children will have had some early education, either in the maintained or PVI sector, before they have an FSP. So any progress assumed between FSP and KS1 could be an overestimate of the true progress made. A weakness of the pupil progress method, namely its lack of a measured baseline 'level of knowledge', is merely replicated with the inclusion of FSP, albeit at an earlier stage.

7.8 An alternative source of measurement for the quality of early years education may be Ofsted inspections. Since 2001, inspectors have been assessing nurseries. The inspections give grades on the 'overall effectiveness' of the school, as well as on specific areas such as teaching and the behaviour of pupils. Taking the example of teaching, Ofsted's 2002/3 Annual Report shows that the percentage of nursery schools attaining a 'good or better' in teaching has increased by 50% since 1994/5¹⁷.

Sixth-form education

7.11 The Atkinson Review recommends the inclusion of post-16 attainment when measuring quality in education. We therefore explore the possibility of extending exam-based methods to cover Sixth form education.

7.12 The main difficulty here is one of data availability: the data available for current and past years are partial, and risks giving an incomplete picture of output in the 6th-form phase. This could lead to a potential discontinuity in official time series, as comprehensive administrative data should be available from 2004 onwards.

7.13 We are exploring whether it is desirable to extend attainment measures to the 6th-form phase for years prior to 2004, or whether it is preferable to assume that output in these years moves in line with the output of compulsory education. There are two data sources currently available for years prior to

¹⁶ The foundation stage profiles are designed to aid teacher assessment, not to provide national statistics; therefore only 10% of the results are centrally collected.

¹⁷ See: <http://www.ofsted.gov.uk/publications/index.cfm?fuseaction=pubs.summary&id=3501>

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

2004, though neither is fully comprehensive:

- School and College Achievement and Attainment Tables (SCAAT) give the numbers of pupils in each school who achieve 2+ A-levels at age 18, making it *feasible* to extend any of the attainment methods we have discussed. The problem is that SCAAT do not include results for pupils registering achievements below A-Level in the 16-18 phase, including for pupils who achieve 5+ A*-C GCSEs for the first time in this phase.
- Based on preliminary analysis of pilot administrative data which exists for 2003, it is clear that there are a significant number of GCSE qualifications attained between 16 and 18 which the SCAAT will not register. In addition they omit certain vocational qualifications; although GNVQs and Vocational A-levels are captured, other vocational qualifications including NVQs are not;
- The Youth Cohort Study (YCS) registers GCSE and A level attainment for 6th form phase and includes of a wide range of vocational qualifications - suggesting a much smaller discontinuity in the time series if YCS data were used for years before 2004 and replaced by administrative data for after that year;
- That discontinuity is less of a problem with the YCS is supported by preliminary analysis which shows a similar level of GCSE attainment in both datasets. The problem is that the data are based on a sample of young people and do not include all individuals, opening the possibility that results may not always be representative of the whole cohort. In addition, the YCS is collected only every two years, making it more difficult to produce estimates for yearly output.

7.14 In spite of these difficulties, YCS is the preferred option for the GCSE and simple cohort progress methods as it can easily be applied. For the cohort progress method, this will involve similar assumptions to those used to map GCSE attainment on to APS attainment for compulsory education.

7.15 For 2004 onwards we expect to have a comprehensive (matched) pupil-level dataset for all pupils in England, including GCSE and A-level qualifications achieved in the 16-18 phase and a fuller range of vocational qualifications. The first estimates of attainment using these data will be available for the cohort who are aged 19 in 2004 (attainment up to and including academic age 18 in 2003/04). As mentioned above, data has been matched for 2002 and 2003 to test the feasibility of the approach; however it is incomplete and cannot be used in published work. The first set of available data was published in February 2005 and we will examine this in more detail to assess its future contribution to this work.

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

Annex 1: Overview of the National Curriculum System

A1.1 Almost all children now have some experience of early years education prior to compulsory education, either state-provided or provided by the private, voluntary or independent sector.

A1.2 Compulsory education for children in England begins from the September after a child turns five and lasts until age 16. It is divided into four National Curriculum Key Stages, with pupils sitting national examinations at the end of each phase. The Key Stages are referred to as Key Stage 1, 2, 3 and 4 and undertaken in numerical order. They vary in length from 2 years to 4 years as detailed below.

A1.3 Pupils are expected to progress by 3 National Curriculum points per year, with the total expected progress for the Key Stages varying according to the length of the phase.

A1.1: Key Stage descriptions				
	Length of Key Stage (age)	Expected Level	Corresponding expected points (at end of KS)	Expected Progress
Key Stage 1	2 years (5-7)	Level 2	15	6 NC points
Key Stage 2	4 years (7-11)	Level 4	27	12 NC points
Key Stage 3	3 years (11-14)	Level 5/6	33/36	9 NC points
GCSE	2 years (14-16)	n/a	n/a	6 NC points

A1.4 The National Pupil Database records the attainment of individual pupils at each KS. By comparing attainment at successive Key Stage examinations, we can measure a pupil's progress.

A1.5 In 2003, around 66% of 17-year-olds stayed on in post-compulsory education, completing 'A' levels or vocational qualifications in school sixth forms or FE colleges¹⁸. Many of these people will go on to enter higher education (HE) to obtain a degree or other qualification.

¹⁸ This includes full-time and part-time students attending maintained schools, independent schools, Sixth Form Colleges, General FE, Tertiary and special colleges and Higher Education Institutions. Source: 'Participation in Education, Training and Employment by 16-18 Year Olds in England: 2002 and 2003', table 2.

Link: <http://www.dfes.gov.uk/rsgateway/DB/SFR/s000469/tab002.shtml>

MEASURING OUTPUT FROM THE EDUCATION SYSTEM

Annex 2: Converting GCSE on to the KS, 2 And 3 APS Scale

A2.1 Pupils in England sit statutory Key Stage 1, 2, 3 and 4 (GCSE) tests at the ages of 7, 11 and 14 and 16 respectively. Key Stage 1, 2 and 3 exams are marked on a common National Curriculum scale, making it possible to measure pupil progress in these phases simply by comparing attainment before and after the Key Stage. However, this comparison is not possible for KS4 as KS4 attainment is measured on a scale which is incomparable with KS1-3. Our response is to create a mapping which expresses GCSE attainment in terms of the National Curriculum point score scale thus allowing us to measure pupil progress between KS3 and KS4.

Proposed method of conversion

A2.2 It is possible to fit a relationship between pupil attainment at KS3 and attainment two years later at GCSE, using regression analysis. To do this, we draw on 2003 matched data, measuring KS3 attainment on the KS1-3 scale and GCSE attainment on the GCSE scale. The fitted relationship is broadly linear as described by the following equation:

$$\text{KS3 attainment on APS scale} = (\text{GCSE attainment} + 44.12) / (2.53)$$

A2.3 On its own, this equation enables us to identify the *KS3 prior attainment* for pupils who have achieved a given level of GCSE attainment two years later at GCSE. However our objective is arrive at an estimate of final attainment for these pupils, although this must be expressed in terms of Key Stage Average Points Score.

A2.4 To achieve this, we begin with the KS3 mark predicted as the prior attainment for a pupil of given final GCSE attainment and add 6 Key Stage points. An assumption of 6 points represents the progress which pupils are expected to make in the 2 years of the KS4 phase. 6 points is the standard assumption because the KS1-3 exam scales were constructed in a way that pupils progress, on average, around 6 points ever 2 years. Thus:

$$\text{KS4 attainment} = [(\text{GCSE} + 44.12) / 2.53] + 6 = (\text{GCSE} + 59.3) / 2.53$$

A2.5 This equation gives the transformation we have applied to convert GCSE attainment, expressed in terms of GCSE Average Points Score, on to the Key Stage Average Points Score scale.

A2.6 However, this transformation involves a statistical restriction which, we suspect, could influence the percentage changes in progress used as the basis of an output index. By using this transformation, we are in effect assuming that KS4 progress is 6 NC points in the base year (2003). Absolute changes in KS4 progress over time will be measured accurately in spite of this assumption, but, when expressed in percentage terms, these changes may be inaccurate, if progress is not equal to 6 NC points in the base year. 2003 must be taken as a “fixed point”, and this could affect the scaling implicit in percentage changes. However we expect this effect will be slight.